УДК 519.651

РЕКУРСИВНЫЙ АЛГОРИТМ ПОСТРОЕНИЯ ФУНКЦИИ ДИРИХЛЕ

А.Н. ЛЫСЮК, С.С. ДЕРЕЧЕННИК

Брестский государственный технический университет
Московская, 267, Брест, 224017, Беларусь

Поступила в редакцию 30 мая 2012

Рассмотрена задача определения общего количества рациональных дробей, значения которых одинаковы и равны x. Продемонстрирована важность данной задачи для процедуры обработки статистических данных, представляющих собой соотношение двух дискретных величин с переменным знаменателем. Установлено, что искомое количество дробей равно значению функции Дирихле в точке x, для построения которой предложено оригинальное порождающее правило, имеющее простую геометрическую интерпретацию. Предложен вариант реализации данного алгоритма, показана его вычислительная эффективность, отмечена его важность в задачах, требующих генерации взаимно простых чисел.

Ключевые слова: функция Дирихле, фрактал, рекурсивный алгоритм, взаимно простые числа, статистическая обработка данных.

Введение

Рассмотрим следующий математический объект. Пусть имеется множество $\mathbb{D}$ рациональных дробей следующего вида:

$$\mathbb{D} = \left\{ \frac{m}{n} : n \in \mathbb{N}, m \in \mathbb{N}_0, m \le n \right\}. \tag{1}$$

Каждая дробь такого множества представляет собой рациональное число, принадлежащее отрезку $[0;1]$. Найдем количество дробей, таких, что их значение одинаково и равно некоторому рациональному числу x из указанного отрезка. Очевидно, что для любого x существует бесконечное количество дробей вида $\frac{c \cdot m}{c \cdot n}$, где $c \in \mathbb{N}$ и $\gcd(n,m) = 1$. Поэтому выполним нормирование искомого количества дробей следующим образом: разделим данное значение на общее количество натуральных чисел. В результате получим некоторую функцию $f(x)$, показанную на рис. 1, где по оси абсцисс располагаются значения x, а по оси ординат нормированное количество дробей из множества $\mathbb{D}$, значение которых равно x. Кроме того, на данном графике также показаны нормали, опущенные из точек функции $f(x)$ на ось абсцисс. Как будет показано далее, данные нормали представляют собой основу для построения фрактального объекта, характеризующего функцию $f(x)$.

В статье рассматриваются основные свойства и способы формирования данной функции, а также показывается ее значимость в статистической обработке данных, представленных соотношением двух дискретных величин.

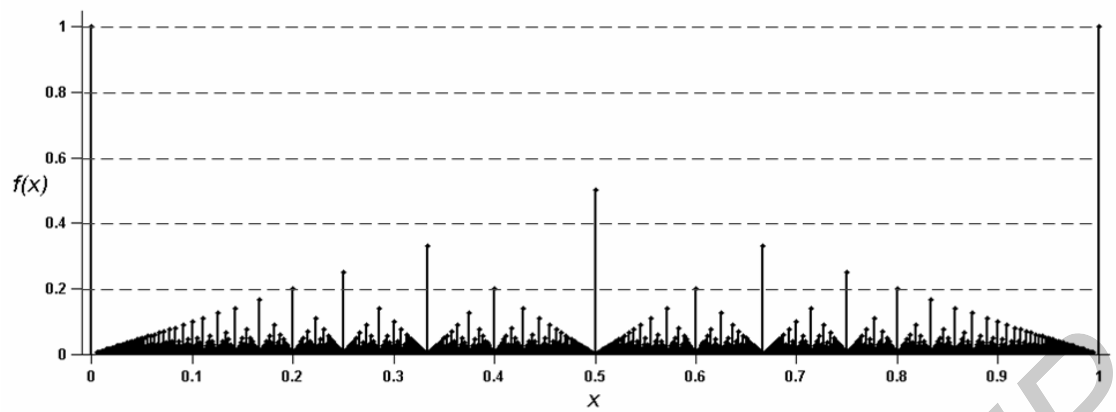


Рис. 1. К задаче определения количества дробей, равных значению x

Статистический анализ соотношения дискретных величин

Предположим, что имеется некоторая генеральная совокупность объектов произвольной природы, в которой выделяется класс «особых» объектов. Примером может служить партия деталей, в которой присутствуют бракованные изделия. Перед аналитиком ставится задача выяснить характер распределения «особых» объектов в генеральной совокупности на основании данных из L выборок. Если для каждой выборки известны значения n_i и m_i, соответствующие общему размеру выборки и количеству особых объектов в ней, то задача может быть решена путем анализа следующего соотношения двух дискретных величин:

$$x_i = \frac{m_i}{n_i}, \quad i = 1, 2, ..., L. \qquad (2)$$

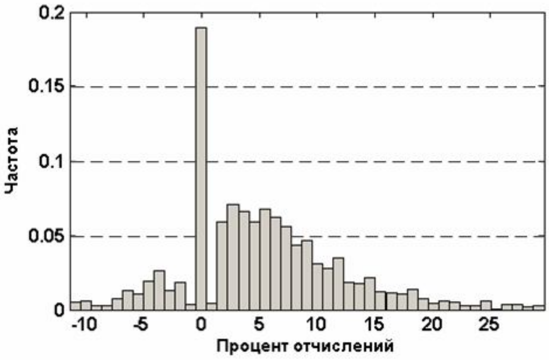
В зависимости от значения величины n_i, следует рассматривать два случая:

Первый случай. Если размеры выборок одинаковы ($n_i = n_j$ для всех выборок), то можно воспользоваться методами классической математической статистики, в результате чего будет построено дискретное эмпирическое распределение, удовлетворяющее решению поставленной задачи.

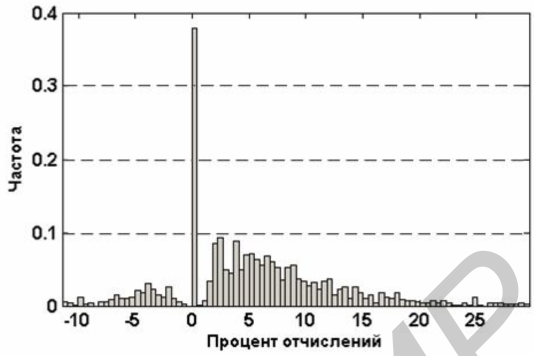
Второй случай. Если размеры выборок неоднородны по численному составу ($n_i \neq n_j$), то возникает ситуация, когда знаменатели в (2) различаются и при построении эмпирического распределения (гистограммы) мы будем сталкиваться с эффектом, показанным на рис. 1. Чтобы понять, почему это происходит, рассмотрим пример, в котором «особые» объекты распределены в выборках равновероятно. В этом случае вероятность того, что величина m_i примет конкретное целое значение из промежутка $[0; n_i]$ равна $1/(n_i + 1)$. Если при этом всем n_i попарно-различны, и значение L стремится к бесконечности, то ряд значений x_i будет соответствовать множеству $\mathbb{D}$ и задача построения теоретического распределения сводится к ранее сформулированной задаче подсчета количества дробей.

Примером реальной задачи является нахождение распределения относительного количества отчисленных/восстановленных студентов в Брестском государственном техническом университете, где в качестве выборок выступали академические потоки, представленные своим размером n_i (численностью студентов) и количеством отчисленных/восстановленных (за учебный год) студентов m_i. В результате обработки выборок были получены гистограммы, приведенные на рис. 2,*а* и 2,*б*, на которых видно, что для величины 0 соответствующая частота максимальна, тогда как для величин, близких к нулю, частоты имеют локальные минимумы.

Нерегулярности формы полученных гистограмм главным образом обусловлены эффектами, упомянутыми во введении и отраженными на рис. 1. При дальнейшей обработке таких гистограмм необходимо использовать специальные методы сглаживания, основанные, например, на свойствах операции округления. Выявление различных, в т.ч. фрактальных особенностей исследуемого математического объекта, может быть полезно с точки зрения повышения эффективности методов статистической обработки данных в аналогичных задачах.



а



б

Рис. 2. Гистограмма процента отчислений, построенная на 45 интервалах (*а*); гистограмма процента отчислений, построенная на 90 интервалах (*б*)

Связь с функцией Дирихле

Как известно, любое неотрицательное рациональное число может быть единственным образом представлено в виде несократимой дроби [1]:

$$x = \frac{m}{n} : x \in \mathbb{Q}, m \in \mathbb{N}_0, n \in \mathbb{N}, \gcd(m,n) = 1\,. \tag{3}$$

Ограничимся рассмотрением рациональных чисел на отрезке $[0;1]$. Как уже отмечалось, для заданного числа x существует бесконечное количество сократимых дробей, принадлежащих множеству $\mathbb{D}$, вида $\frac{c \cdot m}{c \cdot n}$, где $c \in \mathbb{N}$. Для нахождения функции $f(x)$ воспользуемся тем фактом, что значения знаменателя $c \cdot n$ равномерно распределены на множестве натуральных чисел. На основании того, что каждому фиксированному значению знаменателя соответствует единственное значение числителя при заданном x, и так как $n \geq m$, то общее количество дробей, равных x, будет эквивалентно общему количеству различных знаменателей $\frac{|\mathbb{N}|}{n}$, где $|\mathbb{N}|$ – размерность множества натуральных чисел. Поэтому значение $f(x)$ можно рассчитать следующим образом:

$$f(x) = \frac{|\mathbb{N}|}{n} \cdot \frac{1}{|\mathbb{N}|} = \frac{1}{n}. \tag{4}$$

Данная функция будет эквивалентна модифицированной функции Дирихле [2] и может рассматриваться в качестве ее альтернативного определения через количество дробей из множества $\mathbb{D}$:

$$f(x) = \begin{cases} \frac{1}{n} & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}. \tag{5}$$

Традиционный алгоритм построения данной функции заключается в итерационном переборе натуральных значений n и m, при этом если $\gcd(n,m) = 1$, то точка с координатами $\left(\frac{m}{n}, \frac{1}{n}\right)$ будет принадлежать данной функции. Данный алгоритм обладает тем недостатком, что для проверки принадлежности точки функции необходимо выполнять проверку чисел n и m на взаимную простоту, что требует порядка $O\{\lg n\}$ операций [3].

Рассмотрим альтернативный способ нахождения данной функции, основанный на генерации соответствующего фрактала.

Определение фрактала пересекающихся отрезков

Рассмотрим следующее порождающее правило.

<u>Шаг 1:</u> пусть имеются два отрезка нормали, заданные координатами своих вершин: $(x_1, 0)$, (x_1, y_1) для первого отрезка; $(x_2, 0)$, (x_2, y_2) для второго отрезка, причем $x_1 < x_2$ и между этими нормалями на данном шаге нет ни одной промежуточной нормали, т.е. они являются соседними.

<u>Шаг 2:</u> построим еще два отрезка, путем соединения вершин $(x_1, 0)$, (x_2, y_2) и (x_1, y_1), $(x_2, 0)$ соответственно.

<u>Шаг 3:</u> найдем координаты точки пересечения двух, образованных на предыдущем шаге отрезков:

$$x_* = \frac{x_1 \cdot y_2 + x_2 \cdot y_1}{y_1 + y_2}; \;\; y_* = \frac{y_1 \cdot y_2}{y_1 + y_2}. \tag{6}$$

<u>Шаг 4:</u> опустим из найденной точки пересечения нормаль на ось абсцисс, тем самым формируя новый элемент искомой фигуры.

<u>Шаг 5:</u> переходим к шагу 1, выбирая в качестве исходных две соседние нормали.

В качестве начальных используются отрезки-нормали с координатами $(0,0)$, $(0,1)$ для первого отрезка и $(1,0)$, $(1,1)$ – для второго.

Применение данного правила для нескольких циклов показано на рис. 3. Из приведенного рекуррентного способа формирования данной геометрической фигуры можно сделать заключение, что полученный математический объект будет являться фракталом.

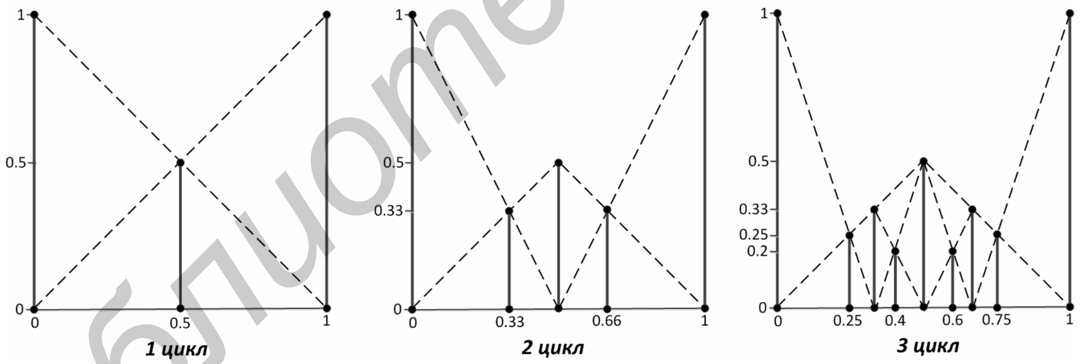


Рис. 3.Фрактал пересекающихся отрезков

На основании (6) можно сделать следующие выводы:

- каждый образуемый отрезок нормали располагается строго между двумя порождающими отрезками, его высота меньше высоты любого из них, но строго больше нуля;
- если вершины порождающих нормалей имеют рациональные значения, то и образуемый отрезок также обладает данным свойством.

Докажем тождественность данного фрактала функции Дирихле. Для чего докажем утверждение *для всех соседних отрезков нормали выполняется следующий ряд свойств:*

$$\begin{cases} x_1 = \dfrac{m_1}{n_1}; y_1 = \dfrac{1}{n_1}; x_2 = \dfrac{m_2}{n_2}; y_2 = \dfrac{1}{n_2}; \\ m_2 \cdot n_1 - m_1 \cdot n_2 = 1. \end{cases} \tag{7}$$

Действительно, предположим, что на некоторой итерации рекурсивного выполнения порождающей функции выполняется данное утверждение, тогда выполним подстановку значений из (7) в (6). В результате получаем:

$$\begin{cases} x_* = \dfrac{m_1 + m_2}{n_1 + n_2} = \dfrac{m_*}{n_*};\ y_* = \dfrac{1}{n_1 + n_2} = \dfrac{1}{n_*}; \\ m_* \cdot n_1 - m_1 \cdot n_* = 1;\ m_2 \cdot n_* - m_* \cdot n_2 = 1. \end{cases} \tag{8}$$

Таким образом, сформулированное утверждение выполняется и для вновь образованного отрезка нормали. Несложно также убедиться в справедливости данного утверждения и для первоначальных отрезков нормали ($n_1 = 1$, $n_2 = 1$, $m_1 = 0$, $m_2 = 1$), на основании чего методом математической индукции можно обобщить полученный результат абсолютно для всех отрезков нормали.

Важно отметить тот факт, что нижние строки уравнений в (7) и (8) удовлетворяют соотношению Безу для взаимно простых чисел, согласно которому наибольший общий делитель чисел *a, b* можно всегда представить как линейную комбинацию *a* и *b* с целыми коэффициентами [4]. Таким образом, пары чисел (m_1, n_1), (m_2, n_2) и (m_*, n_*) – взаимно простые. Следовательно, рассматриваемый фрактал приводит к формированию функции Дирихле и также может служить одним из ее представлений. Формируемые таким образом дроби (см. формулы 8) носят название дробей Фарея, открытых им в 1816 году [5].

Порождающая функция, рассматриваемая выше, может выступать в качестве алгоритма вычисления функции Дирихле. При этом целесообразнее использовать соотношения (8), а не (6). В этом случае для работы алгоритма необходимы два целочисленных массива, хранящие значения m_i и n_i. В момент инициализации данные массивы будут хранить значения $\{0;1\}$ и $\{1;1\}$, соответственно. Во время выполнения очередного цикла для каждого массива выполняются два действия: суммируются значения в двух соседних ячейках, и результат суммирования помещается в ячейку массива между этими двумя ячейками. Иллюстрация данного алгоритма представлена на рис. 4, на котором элементы, формируемые на очередном этапе, показаны в прямоугольнике.

	Массив m	Массив n
Цикл 0	0 1	1 1
Цикл 1	0 $\boxed{1}$ 1	1 $\boxed{2}$ 1
Цикл 2	0 $\boxed{1}$ 1 $\boxed{2}$ 1	1 $\boxed{3}$ 2 $\boxed{3}$ 1
Цикл 3	0 $\boxed{1}$ 1 $\boxed{2}$ 1 $\boxed{3}$ 2 $\boxed{3}$ 1	1 $\boxed{4}$ 3 $\boxed{5}$ 2 $\boxed{5}$ 3 $\boxed{4}$ 1
Цикл 4	0 $\boxed{1}$ 1 $\boxed{2}$ 1 $\boxed{3}$ 2 $\boxed{3}$ 1 $\boxed{4}$ 3 $\boxed{5}$ 2 $\boxed{5}$ 3 $\boxed{4}$ 1	1 $\boxed{5}$ 4 $\boxed{7}$ 3 $\boxed{8}$ 5 $\boxed{7}$ 2 $\boxed{7}$ 5 $\boxed{8}$ 3 $\boxed{7}$ 4 $\boxed{5}$ 1
…	…	…

Рис. 4. Иллюстрация алгоритма построения фрактала

Преимущество данного алгоритма очевидно, так как новые значения точек функции Дирихле получаются за $O\{1\}$, а не за $O\{\lg n\}$ операций. Отметим еще одну важную особенность предложенного алгоритма. Так, в результате его работы образуются пары взаимно простых чисел, соответствующие элементам рабочих массивов с одинаковыми индексами. Следовательно, алгоритм может использоваться в качестве генератора взаимно простых чисел, что, например, может быть востребовано в системе шифрования Хилла [6] или системе подстановки Цезаря [7]. Кроме того, взаимно простые числа применимы при построении простейших линейных конгруэнтных генераторов псевдослучайных чисел вида $X_{k+1} = (X_k + m_i) \bmod n_i$ с периодом равным n_i [8].

Заключение

В данной работе была рассмотрена задача определения общего количества рациональных дробей из множества $\mathbb{D}$, значения которых одинаковы и равны x. Продемонстрирована важность данной задачи для процедуры обработки статистических данных, представляющих собой соотношение двух дискретных величин с переменным знаменателем. Отмечено, что возникающий при этом эффект приводит к заметному искажению результирующей гистограммы, что затрудняет ее восприятие и интерпретацию и требует дополнительной обработки.

В ходе исследования установлено, что искомое количество дробей равно значению функции Дирихле в точке x, для построения которой предложено оригинальное порождающее правило, имеющее простую геометрическую интерпретацию. Тем самым был получен фрактальный объект, соответствующий функции Дирихле и имеющий, в отличие от нее, более простой алгоритм формирования. Предложен вариант реализации данного алгоритма, показана его вычислительная эффективность. Алгоритм может применяться, в том числе, в задачах, требующих генерации взаимно простых чисел.

THE RECURSIVE ALGORITHM TO CONSTRUCT DIRICHLET FUNCTION

A.N. LYSIUK, S.S. DERECHENNIK

Abstract

The problem of determining the total number of rational fractions with values are equal to *x* is considered. The importance of this problem is demonstrated for the procedures processing statistical data, representing the ratio of two discrete variables with a variable denominator. It is found out, that that required number of fractions is equal to the Dirichlet function value at the *x* point, and the original rule is proposed to construct it, which has a simple geometric interpretation. A proposed implementation of this algorithm shows its computational efficiency, and its importance is noted for problems requiring generation of relatively prime numbers.

Список литературы

1. *Александров П.С.* Введение в теорию множеств и общую топологию. М., 1977.
2. *Bruckner A., Bruckner J., Thomson B.* Elementary Real Analysis. NJ, 2008.
3. *Кнут Д.Э.* Искусство программирования. Получисленные методы. М., 2007.
4. *Виноградов И.М.* Основы теории чисел. М., 1965.
5. *Apostol T.M.* Modular Functions and Dirichlet Series in Number Theory. NY, 1997.
6. *Hill L.S.* // The American Mathematical Monthly. 1929. Vol. 36. P. 306-312.
7. *Kahn D.* The Codebreakers: The Story of Secret Writing. NY, 1996.
8. *Gentle J.* Random Number Generation and Monte Carlo Methods. NY, 2003.